

Towards a Morphosyntactic Account of Taiwanese Question Particle *kam*

Chu-Ren Huang

This article accounts for the interrogative particle *kam* of the South Min dialect spoken on Taiwan as a second-position sentential clitic in GPSG (Generalized Phrase Structure Grammar) and discusses the theoretical implications of such an account. Because of its free selection of host words, its lack of morphological idiosyncrasies, the inapplicability of syntactic rules to it and its closure of affixation, *kam* cannot be an affix. Furthermore, based on its prosodic domain, binding, replacement, ordering, and distribution features, *kam* is shown not to be a word. Following the theory of cliticization outlined in Zwicky and Pullum (1983) and Zwicky (1985), *kam* is proved to be a clitic.

A formal account of the cliticization of *kam* in GPSG is given, adopting the notion of Immediate Precedence (IP) proposed in Nevis (1986) and Zwicky and Nevis (1986). A formal definition of IP is proposed within the GPSG theory. The possible interaction between IP and LP (Linear Precedence) with existential implication as proposed in Huang (1985) is also discussed. Finally, it is suggested that the existence of *kam* as a second-position proclitic in a language dominated by sentence final enclitics is worth in-depth future studies.

I. Introduction

Huang (1985) argues that Mandarin sentential question particles are sentential clitics occurring in the absolute sentence-final position. This position is followed and strengthened in Sheu (1987). Shiu and Huang (1988) elaborate this analysis by giving it a formal account in LFG. It is also observed in Huang (1985) that the location of Chinese sentential clitics has several interesting theoretical implications. Previously, Zwicky (1977) observed that the sentence-final position is the least common position for

clitics. And Kaisse (1982) proposed as a universal rule that sentential clitics can only occur in the second position, following Wackernagel's Law. For the claim made by Zwicky, the Chinese data pose some difficulties, even though one exception does not necessarily nullify generalizations of markedness. However, for Kaisse's universal, given in (1), the absolute sentence-final position for Chinese sentential clitics is one indisputable counterexample.

(1) Sentential Clitics may only occur in second position.

(Kaisse 1982:12)

Typologically speaking, one explanation may be that the second position is the unmarked position for cliticization in well-studied Indo-European languages and that the sentence-final position is the unmarked position for cliticization in Chinese or even Sino-Tibetan languages.¹ To carry this hypothesis to the extreme, one might hypothesize that positions of cliticization are simple typologically set (i. e. as a parameter set for languages families in Universal Grammar), and thus explains Kaisse's (1982) postulation as influenced by Indo-European and other data. In trying to build up an adequate theory of cliticization and trying to give a formal and rigorous grammar of Chinese languages, one cannot overlook the possibility that positions of cliticization (second-position vs. sentence-final) are typologically determined and, regardless of whether the hypothesis is correct or not, how do the two positions of cliticization interact with each other in these languages.

Zhu (1985) observes that, in addition to the more often seen form of A-not-A questions, there are Chinese languages forming Yes-no questions

1 I am not aware of any in-depth theoretical study of cliticization of Tibetan languages. However, Sun (1988 & personal communication) supplied data from a couple of Tibetan languages with particles which seem to be either sentence-final enclitics or proclitics. See the preface to Huang (1988) for a sketch of some of the data.

with a so-called 'interrogative adverb' attached before a VP.² The languages adopting this marked form are the Wu dialects, South-western Mandarin and Lower Yang-Tze Mandarin.³ The set of pre-VP positions overlap, if not coincide with the set of second positions for cliticization. An identical syntactic pattern is followed in another Chinese Language--South Min. This paper studies Taiwanese, a South-Min dialect spoken on Taiwan. It will be argued that the pre-predicate question particle *kam* in Taiwanese is indeed a second position clitic. A formal account in GPSG (Generalized Phrase Structure Grammar, Gazdar et al. 1985) will also be given to this clitic.

II. The Theory of Cliticization

Clitics and particles are two traditional terms used to refer to morphemes whose grammatical status is unclear. Due to the broad pre-theoretical uses of the term, linguistic theories do not offer a clear-cut definition of clitics until the progress of recent theoretical studies of them. Zwicky (1977) presented a fairly thorough theoretical survey of clitics, and Klavans (1982) proposed the definition that clitics are phrasal affixes. Even though Klavans's definition and her parameters are not universally accepted, the concept that clitics are attached to phrases instead of morphemes is accepted as the general characterization of cliticization.⁴ Zwicky and Pullum (1988) adopted

2 Zhu (1985) used, in Chinese, the term 'Chinese dialects' instead of 'Chinese languages.' The choice of wording is influenced by cultural considerations which I will not elaborate on here.

3 I was not aware of the occurrence of this form in Chinese dialects other than Taiwanese and Shanghainese, a Wu dialect. I am indebted to Laurent Sargat for referring me to Zhu's article and for pointing out the fact that this occurs in other Chinese languages.

4 More specifically, Nevis (1986) proposes to classify clitics into two subcategories unified by the characteristics of phonological liaison, one of the phrasal affix type and one of the type of bound words. The second type refers to a group of clitics whose occurrences are constrained but still possess wordlike properties.

a similar position to argue for a separate module for cliticization. Since clitics are attached to phrasal categories, cliticization must utilize information from the syntax module, thus the module of cliticization cannot precede the module of syntax. In addition, it is shown that no syntactic rule can affect a cliticized element, hence cliticization must follow syntax.

Klavans (1982) proposed that clitics be characterized with three parameters: P1 Initial/Final, P2 Before/After and P3 Proclitic/Enclitic. The first parameter determines the domain (a phrasal host) and the position (at the front edge or the final edge of the phrase) of cliticization. The second parameter determines the placement (before or after the edge specified by P1) of a clitic. The last parameter determines the direction of phonological liaison (a clitic leans to the word preceding it or following it). P2 and P3 are independently motivated because it is observed that the direction of syntactic attachment does not necessarily coincide with the direction of phonological liaison.⁵ In Klavans (1985), she replaces the names of the parameters with Immediate Dominance, Linear Precedence and Liaison respectively. Nevis (1986) observes that the first two parameters have functions

5 Klavans (1982) cites Ngunhcarara data to support this position. A more familiar case is English *be* verbs. Though forming a predicative phrase with the following constituent, the clitic counterparts of *be* verbs, such as *'s* and *'m*, phonologically lean on the previous words, such as *the man next to you's* and *I'm*. In Mandarin Chinese, I have shown (Huang 1985 and 1987) that the NP clitic *de* is placed before the last constituent of an NP (i. e. the head) but phonologically attached to the word preceding it. Thus Mandarin Chinese offers another strong case for treating syntactic and phonological attachment of clitics separately. The *de* data are given in (i). I am following the convention in the field to use '=' to mark phonological liaison in cliticization.

(i) a. hung=de shu
 red DE_{NP} book
 'red book (s)'
 b. Libai shie=de shu
 Li-Po write=DE_{NP} book
 'the book written by Li-Po'

exactly like the ID rules and LP statements in GPSG. Thus he proposes that Liaison be the only parameter specific to cliticization by arguing that mechanisms to treat phenomena relevant to the first two parameters are independently motivated in syntax and the postulation of them is unnecessary. Nevis (1986) and Zwicky and Nevis (1986) offered formal accounts of cliticization following this line of argument in GPSG. I will by and large follow their assumptions and formalism.

To solve the traditionally fuzzy definition of clitics, Zwicky and Pullum (1983) and Zwicky (1985) established the standard procedure for identifying a clitic. Clitics are hard to characterize because they are both word-like and affix-like, the best way to determine its identity is to show that it is neither a word nor an affix. Zwicky and Pullum (1983) showed that clitics are unlike affixes in that they select their phrasal hosts rather than lexical hosts. And Zwicky (1985) showed that clitics are not like words in that they are usually phonologically dependent on their hosts and that they cannot occur in isolation. In other words, the identities of clitics are determined by elimination, by showing that they are neither words nor affixes. I will use criteria from both articles, with adaptations shown to be useful for Chinese languages in Huang (1985) and in Sheu (1987), to show that Taiwanese *kam* is indeed a sentential clitic.

III. The Data

Cheng (1977) describes the predicate-initial *kam* as the unmarked form of Taiwanese question particle. As a matter of fact, *kam* is the only non-sentence-final question particle in this language. The following data show that *kam* is a candidate for second position clitics.⁶

6 The romanization of Taiwanese in this article follows Cheng and Cheng (1977) except for the substitution of *ts* for *ch* and for the omission of diacritic tone markers.

- (2) a. i kam u lai?
s/he KAM PERF come
'Did s/he come?'
b. i kam bat li?
s/he KAM know you
'Does s/he know you?'
c. goa m tsaiiã i kam bat li
I NEG know s/he KAM know you
'I don't know whether he knows you or not.'

The three sentences in (2) show that *kam* occurs before an intransitive verb phrase (2a), a transitive verb phrase (2b), and in an embedded sentence (2c). Hence it is tempting to postulate *kam* as a verbal prefix since it occurs immediately before a verb and seems to be attached to the verb in all three cases. However, *kam* does occur in positions non-concatenating to a verb, such as in (3).

- (3) a. i kam tsang lai (e)
s/he KAM yesterday come E
'Was it yesterday that s/he came?'
b. i kam tiãtiã lai?
s/he KAM often come
'Does s/he come often?'
c. tsang kam i die kuankun
yesterday KAM s/he get champion
'Was is it s/he who won the championship yesterday?'

Based on sentences in (3), it could be concluded that the hypothesis that *kam* might be a verbal prefix is incorrect. This Taiwanese question particle will be examined and shown to be a sentential clitic.

IV. Kam as a Sentential Clitic

Data such as (2) and (3) suggest that the crucial step in establishing the

cliticness of *kam* is to clarify that it is not an affix. I will show that it is unlike an affix based on the free selection of host words, the lack of morphological idiosyncrasies, the inapplicability of syntactic rules and the closure of affixation.^{7, 8}

First, both inflectional and derivational affixes are known to select the grammatical categories of the stems they are attached to. For instance, Mandarin plural suffix *-men* can only be attached to a human noun stem. Clitics, on the other hand, select their phrasal hosts instead of their host words. Thus, unlike affixes, clitics impose no restriction on the categories of the host words they are attached to. (2) and (3) demonstrate clearly that *kam* can be attached before a verb, an adverbial and a noun. Therefore *kam* does not seem to be a verbal prefix in particular nor an affix in general.

Second, morphological affixation may show idiosyncrasies, such as in the suppletive form of the English past tense in the verb *caught* (as opposed to the regular past tense affix *-ed* in *joined*). Taiwanese *kam* does not vary its phonological shape and is not likely to be an affix.

Third, affixes form syntactic words with their host stems, therefore they form natural constituents with regard to syntactic operation. On the other hand, since we take cliticization as phrasal affixation, a clitic is in construction with a whole phrase and does not form a syntactic unit with its host word (the word it is phonologically attached to). Thus it is predicted that syntactic rules affect an affix together with its host phrase as

7 Take note that, following Klavans (1982), I distinguish host words from host phrases for cliticization. Since clitics are in general phrasal affixation, the term host, if used alone, refers to the host phrase a clitic is syntactically attached to. The term host word refers to the word which a clitic leans on, often with phonological liaison.

8 There are six criteria to distinguish clitics from affixes proposed in Zwicky and Pullum (1983). Four are adopted here because the other two, when applied to *kam*, do not yield clearly interpretable results.

a unit while they do not affect a clitic together with its host word. This is illustrated in (4), where *kam* cannot undergo topicalization with its immediately following host word.

(4) a. (= 3a)

i kam tsang lai (e)

s/he KAM yesterday come E

'Was it yesterday that s/he came?'

b. * kam tsang i lai (e)

KAM yesterday s/he come E

Fourth, clitics as phrasal affixes often mark the boundary of a phrase. In such cases, morphological affixes cannot cross phrasal boundaries and therefore are closed off by clitics. In other words, clitics can be attached to affixes while affixes cannot be attached outside of clitics. In Taiwanese, *m-* is a verbal prefix of negative polarity.

(5) a. i kam m bat li?

s/he KAM NEG know you

'Doesn't s/he know you?'

b. *i m kam bat li?

s/he NEG KAM know you

Even though both *m-* and *kam* are attached before a verb, *m-* cannot occur outside of *kam*. Since we already know that *m-* is a verbal prefix and that it is necessarily attached to a verb, the closure property indicates that *kam* is a clitic rather than an affix. Thus, all the four applicable criteria indicate that *kam* is not an affix and is very likely a clitic.

The other half of the theoretical proof of the clitic-hood of *kam* consists of the demonstration that it cannot be a word. I will follow the criteria set up in Zwicky (1985) to show that *kam* is not a word. The criteria adopted here include prosodic domain, binding, replacement, ordering, and distribution.⁹

First, like many other tonal languages, the phonological sandhi rules of

Taiwanese apply to words. More specifically, sandhi rules governing the tone value of a syllable are conditioned by the immediate following syllable belonging to the same word. That is, words are the domains of sandhi rules in Taiwanese. It is observed that *kam* is always pronounced with the sandhi 44 high-level (陰平 *yin-ping*) tone, while the lexically assigned rising (上 *shang*) tone of 53 would never surface.¹⁰ In other words, *kam* is prosodically dependent on the following syllables. Given the fact that *kam* falls in the prosodic domain of the sandhi rules without exception, it cannot be an independent word and can only be a clitic.¹¹

Second, words can occur in absolute isolation while bound elements, such as clitics, cannot. *kam* is bound and cannot occur in isolation. This is confirmed by the fact that it only appears in its sandhi form in natural speech, thus it is not likely to be a word. Although binding cannot be the sole criterion in testing clitic-hood because most so-called 'function words' tend not to be able to occur in isolation, binding together with prosodic domain, offer strong support for the classification of *kam* as a clitic.

Third, full-fledged words often have their pro-forms, such as the well-known case of pronouns and proverbs (such as English *do*). In Taiwanese we have *anni* 'in-this-way' and *anzua* 'how' as pro-adverbs,¹² as well as

9 It is worth pointing out again that Zwicky's (1985) list contains all possible criteria for distinguishing clitics from words, regardless of whether these criteria are absolute principles or whether these criteria are applicable in all languages. Hence I am only using the criteria which yield meaningful results in Taiwanese.

10 Take note that, following the convention in Chinese Phonology, the tone names given are the traditional Chinese ones, which are originally descriptive terms but do not reflect the modern tone values any more. Hence a rising tone in Taiwanese has the descending 53 value.

11 I am indebted to Drs. Ho Dah-an and Yang Siou-fang for bringing up the sandhi fact and its significance to the clitic-hood of *kam*.

12 *anni* is also a pro-predicate, such as in *li anni kam ho you-such-KAM-o*. k. 'Is it o.k. for you to do so?'

various pronouns such as *i* 's/he,' *li* 'you (singular)' and *lin* 'you (plural).' There is simply no proform for *kam*. This feature of *kam* is not the same as that of a word.

Fourth, alternations of ordering of words are common. In languages with free word order, ordering changes often signal stylistic changes. In languages with stricter word order, changes in word order are, of course, syntactically significant. On the other hand, order alternation of attached morphemes, clitics or affixes, owing to their bound feature, is not allowed. *Kam* always precedes the predicate of a sentence and never alternates with anything inside the predicate. The ordering behavior of *kam* is not like a word.

Fifth, related to ordering, the distribution of words cannot be characterized with a single principle (a Mandarin Chinese verb occurs in an NP, in a VP or in an adjunct, etc.). However, the distribution of non-words, such as affixes, can often be stated in one simple principle (e.g. the plural affix *-men* of Mandarin occurs right after a noun stem). For instance, see (6).

- (6) a. tsang kam i lai e
yesterday KAM s/he come E
'Was it s/he who came yesterday?'
b. i kam tsang lai e
s/he KAM yesterday come E
'Was it yesterday that s/he came?'

The pair of sentences suggest that the position of *kam* could be captured by the simple principle of either 'second position' or 'immediately before a predicate.' Although a formal account will be delayed until the next section, we can conclude here that the distribution of *kam* is not similar to that of a word. The five applicable criteria unequivocally show that *kam* is not a word.

The above discussion clearly demonstrates that *kam* is neither an affix nor a word, although it has features similar to both words and affixes. Thus, based on the theories of cliticization, we conclude that it is a clitic.

V. A GPSG Account of the Cliticization of *kam*

A formal account of *kam* within the framework of GPSG following Zwicky and Nevis(1986) will be given in this section. First of all, the domain of cliticization has to be determined. It is an uncontroversial fact that *kam* is a marker of interrogation, a discoursal function of a sentence. Thus it semantically functions at the sentential level. Syntactically, we can show that it occurs outside a matrix VP, such as in (7).

- (7) a. i kam bat lai
 s/he KAM know come
 'Has s/he ever come?'
 b. *i bat kam lai
 s/he know KAM come

In (7), *bat* is an (auxiliary) verb which puts the sentence in the experiential aspect.¹³ *bat* clearly marks the boundary of the matrix VP.¹⁴ The ungrammatical(7b) suggests that *kam* lies outside of the matrix VP and can only be dominated by the S node, reflecting its semantic function. Hence the data

13 In many recent syntactic theories, such as GPSG (Gazdar et al. 1985) and LFG (Bresnan 1982), English auxiliaries such as *have* and *be* are treated like raising verbs. In other words, these so-called auxiliaries require two syntactic arguments, an NP and a predicative phrase. However, thematically they require only one argument, the predicate, with the additional stipulation that the syntactic subject of these verbs controls the predicate. I think the semantically bleached verb *bat* (literally 'to know'), used as a marker of the experiential aspect, can be treated along the same line. A detailed analysis, however, is beyond the scope of this short paper.

14 We may also cite as supporting evidence the following pair of sentences. (i) a is taken from Cheng (1977). In both sentences, *ehiau* 'can (ability)' is a modal auxiliary while *eng ti* 'use chopsticks' is an indisputable VP. *kam* can only occur in the sentential second position.

- (i) a. i kam ehiau eng ti
 s/he KAM 'can use chopsticks
 'Can s/he use chopsticks?'
 b. *i ehiau kam eng ti
 s/he can KAM use chopsticks

might be accounted for with the following ID rule.¹⁵

- (8) $S \longrightarrow NP, VP, kam$

However, the above data alone cannot rule out the following alternative analysis, representing the intuitive description of 'predicate-initial'.

- (9) $VP \longrightarrow VP, kam$

More specifically, although (9) would overgenerate the ungrammatical (7b) under our treatment of *bat* 'to know' as a verb subcategorized for a VP, the problem could conceivably be circumvented by assigning *bat* a distinguished category such as AUX. However, the following sentences suggest otherwise.¹⁶

- (10) a. *tsang kam i lai e*
[=7a] yesterday KAM s/he come E
'Was it s/he who came yesterday?'

- b. **kam tsang i lai e*
KAM yesterday s/he come E

- (11) a. **kam i tsang lai e*
KAM s/he yesterday come E

- b. *i kam tsang lai e*
[=7b] s/he KAM yesterday come E
'Was it yesterday that s/he came?'

None of the sentences in either (10) or (11) involve controlled VPs or embedded Ss. Hence *i tsang lai* 's/he came yesterday' in (10a) cannot be a constituent of a VP. Thus (10a) clearly shows that *kam* lies outside a VP and that (9) is not a possible ID rule for the cliticization of *kam*. The

15 Take note that in GPSG the linear precedence relations are accounted for in separate LP statements. Thus the ID rules introduced here do not represent any ordering relationship among the constituents. Notation-wise, the commas separating daughters indicate that there is no precedence relationship implied.

16 It should be clear from the free translation that *kam* also put focus on the constituent immediately following it. I suspect that this foregrounding effect is best captured by a feature, but will leave a detailed account for later studies.

contrast between (10) and (11) also brings out vividly the requirement that *kam* be placed in the second position. The linear order between the subject *i* 's/he' and the adverbial *tsang* 'yesterday' does not affect the grammaticality of Taiwanese sentences since the counterparts of (10) and (11) without *kam* are all grammatical. Thus we can infer that (10b) and (11a) are ruled out due to *kam*'s not being in a second position. We may even conjecture that the alternation of word order between (10a) and (11b) is necessitated by the sentential second position requirement. The above discussion rules out the VP attachment account of *kam*.

An absolute second position, however, does seem to be too strong a requirement. Consider the following examples:

- (12) a. *tsit kiã taitsi, li kam ham i kong koe*
 this item affair you KAM with s/he tell PERF
 'This issue, have you brought it up with him/her?'
 b. *li kam ham i kong koe tsit kiã taitsi*
 you KAM with s/he tell PERF this item affair
 'Have you brought this issue up with him/her?'
 c. *ham i, li kam kong koe tsit kiã taitsi*
 with s/he you KAM tell PERF this item affair
 'With him/her, have you brought this issue up?'

In (12a), *kam* occurs after two phrasal categories, the topic NP *tsit kia taitsi* 'this issue' and the subject NP *li* 'you'. Since *kam* occurs after a subject, it cannot be a daughter of S-bar. Compared with (12b) and (12c), we find that the second position requirement holds within a matrix S.¹⁷ Thus the

17 The sentence (i) is an apparent counterexample. In (i) *kam* seems to occur after two constituents, an NP and a PP. However, we could also take *li ham i* 'you and s/he' as a coordinated NP and explain this apparent exception away.

(i) *li ham i kam kong koe tsit kiã taitsi*
 you with s/he KAM tell PERF this item affair
 'Did you bring this issue up with him/her?'

data in (12) both confirm the ID rule (8) which stipulates that *kam* is dominated by an S, and the second position description. To capture the generalization that *kam* is dominated by a matrix S, we propose the following feature-based ID rule.

(13) $[+MTRX] \longrightarrow XP, H, \{[SUBCAT\ kam]\}$

In (13), like in recent GPSG theory (Gazdar et al. 1985), the feature SUBCAT is used as terminal symbol features for function words. For instance, conjunctions, such as *and* and *or*, and complementizers, such as *that* and *for*, are represented by $\{[SUBCAT\ \beta]\}$, where an additional rule of $\{[SUBCAT\ \beta]\} \longrightarrow \beta$ instantiates the phonological representation. It is observed in Huang (1985 & 1987) that the notion of a matrix sentence figures crucially in the cliticization of the Mandarin sentential enclitic *ma* and *ne*. Thus the non-percolating feature MTRX matrix sentence is independently motivated.¹⁸

Next, we will turn to the precedence relations of the clitic *kam*. Very few constituents occur pre-verbally. In addition to subjects, there are modal adverbs (Cheng 1977), temporal expressions and locative expressions. We will show that regardless of the identity and the number of constituents occurring before the matrix predicate, *kam* always occurs immediately after the first constituent.

(14) Locative PP

- a. gina kam ti taibak thaktseh
 children KAM at Taipei read-book
 'Are the children studying in Taipei?'
- b. *gina ti taibak kam thaktseh
 children at Taipei KAM read-book

18 We assume the following FCR (i) such that it is not necessary to indicate in the ID rule (13) that the mother is an S (i. e. $[+V, -N, BAR\ 2, +SUBJ, -COMP]$). However, it is still not clear to me at this time how to ensure that this feature instantiates on the matrix but not on any embedded sentences.
 (i) $[+MTRX] \supset [+V, -N, BAR\ 2, +SUBJ, -COMP]$

- c. ti taibak kam gina thaktseh
at Taipei KAM children read-book
'In Taipei, is it children who study?'
- d. *ti taibak gina kam thaktseh
at Taipei children KAM read-book

(15) Modal Adverb and Temporal Expression

- a. i kam iting minatsai lai
s/he KAM certainly tomorrow come
'Is it for sure that s/he will come tomorrow?'
- b. *i iting kam minatsai lai
s/he certainly KAM tomorrow come
- c. i kam minatsai iting lai
s/he KAM tomorrow certainly come
'Is it tomorrow that s/he will come for sure?'

(16) Modal Adverb and Locative Expression

- a. i kam iting ti miko thaktseh
s/he KAM certainly at U.S. read-book
'Is it for sure that s/he will come tomorrow?'
- b. *i iting kam ti miko thaktseh
s/he certainly KAM at U.S. read-book

The three sets of data illustrate unambiguously that *kam* occurs in the second position in a sentence, regardless of the number of pre-verbal constituents.¹⁹

19 The absence of a (16c) as a counterpart of (15c) can be easily explained. (i), corresponding to (15c) without *kam*-cliticization, is grammatical, while (ii) is not. The linear precedence relations among pre-verbal adverbials, however, are beyond the scope of this current paper.

(i) i minatsai iting lai
s/he tomorrow certainly come
'S/He will certainly come tomorrow.'

(ii) *i ti miko iting thaktseh
s/he at U. S. certainly read-book

Thus our account of this clitic will have to capture this generalization.

To capture the above generalization, we adopt the notation of the special precedence relation 'immediate precedence (IP)' proposed in Zwicky and Nevis (1986). The immediate precedence relation is defined such that the category to the right of the double 'less-than' sign '<<' is required to immediately follow the category to the left of it. The rule (17) stipulates that the sentential clitic *kam* occurs immediately after its first sister constituent. Since *kam* is dominated by an S, this LP statement in effect guarantees that *kam* occurs at a sentential second position.

(17) [$+$ FIRST] << [SUBCAT *kam*]

The rule (17), however, does not tell us how the first sister is identified. The MARGIN feature FIRST, as proposed in Nevis (1985), is instantiated on exactly one category among all sisters in a local tree. The linear precedence statement (18) will ensure that the category with this feature precedes all its sisters. Given the definition of LP statement in GPSG, it is easy to infer that no more than one sister can bear this feature because two sister constituents with the feature [$+$ FIRST] would violate (18). It is also worth noting that the instantiation of the feature [$+$ FIRST] cannot lead to violation of other LP statements.

(18) [$+$ FIRST] < [$-$ FIRST]

One last loose end that needs to be tied up is to guarantee the appearance of the feature [$+$ FIRST]. Since I am assuming that [$+$ FIRST] is a freely instantiated non-percolating feature, there is nothing in the grammar requiring its instantiation. One possibility is to combine immediate precedence with LP statement with existential implication, first proposed in Huang (1986) for coordination in Mandarin Chinese, to guarantee the co-occurrence of the feature [$+$ FIRST] with [SUBCAT *kam*]. In short, the precedence relation 'e<<' requires that the category on the right hand side immediately follows the category on the left hand side and that the category on the left hand side occurs whenever the category on the right does. The definition of

the acceptability condition of an instance of an entailment-annotated immediate precedence relation (EAIP-acceptable) is given in (19).

(19) $e^{<<}$ -acceptable

Let Φr be the set of projections induced by a rule r , where $r = C_0 \rightarrow C_1, \dots, C_n$, and $e^{<<}$ is the relation of immediate precedence (with existential entailment annotated) determined by the grammar, then any projection ϕ in Φr is $e^{<<}$ -Acceptable (or EAIP-acceptable) if and only if

- (i) For every category G_i, G_j , and G_k which $\phi(C_i), \phi(C_j)$, and $\phi(C_k)$ extends respectively, where $0 < j < n$ and $j \neq k$, for each category G_j and for every Immediate Precedence (IP) Statement $G_j \ll G_i$: ' $\phi(C_k) < \phi(C_i)$ ' \rightarrow ' $\phi(C_k) < \phi(C_j)$.'
- (ii) for every category $\phi(C_i)$, for every category C_i' which the category $\phi(C_i)$ extends, there is no category $C_j, j > n$ (i.e., C_j is not specified in the rule r), such that $C_j e^{<<} C_i'$ is an EAIP-statement.

In prose, the formal definition in (19) stipulates that a local tree is EAIP-acceptable if and only if 1) it is IP-acceptable, and 2) the right-hand side category of an EAIP-statement does not occur unless the left-hand side category also occurs as its sister. The first part of the definition can be extracted as a formal definition of IP-acceptable. It requires that for a local tree to be IP-acceptable, all the categories that linearly precede a certain category, other than the one specified in the immediate precedence relation, must also precede its immediately preceding category. This, of course, guarantees that there is only one immediately preceding category. We will show later that this definition helps capture the generalization of the cliticization of *kam*.

First, the introduction of the more complicated entailment annotated immediate precedence relation on top of IP will be justified by re-examining the data motivating immediate precedence in Nevis (1985) and Zwicky and

Nevis (1986). All linguistic data requiring IP statements involve second position clitics. If acceptability of immediate precedence is defined along the same line as LP-acceptable in Gazdar et al. (1985), it would predict that unless both categories specified in a precedence statement are extended by separate categories on a local tree, the IP statement will be trivially satisfied.²⁰ Since our grammar does not require the MARGIN feature $\llbracket +\text{FIRST} \rrbracket$ to be instantiated on every local tree, it would allow second position clitics to be vacuously constrained by the IP statements and be free to occur in any position when the relevant MARGIN feature is not instantiated. One particular case would be when the feature $\llbracket +\text{FIRST} \rrbracket$ is instantiated on a $\llbracket \text{SUBCAT } kam \rrbracket$ category. Since no categories preceded by the $\llbracket \text{SUBCAT } kam \rrbracket$ category extend the $\llbracket +\text{FIRST} \rrbracket$ category, the local tree would be acceptable with regard to the IP statement (17). Thus the definition would wrongly predict that a second position clitic could occur sentence-initially. Furthermore, any attempt to assign the feature $\llbracket +\text{FIRST} \rrbracket$ to the category linearly preceding all other categories, and thus avoiding wrongly instantiating the feature on $\llbracket \text{SUBCAT } kam \rrbracket$, would create a vicious circle. In other words, without existential entailments, IP statements simply fail to account for the data they are motivated to account for.

Another observation is that the edge-marking feature $\llbracket +\text{FIRST} \rrbracket$ is grammatically relevant only when it co-occurs as a sister of a second position clitic, such as *kam*. This fact is not captured in either Nevis (1985) or Zwicky and Nevis (1986). Nor is the conditions on the instantiation of the MARGIN features clearly articulated. We would argue that both pro-

20 The definition of LP-Acceptable given in Gazdar et al. (1985: 99) is cited here for reference:

Let Φ_r be the set of projections induced by a rule r , where $r = C_0 \longrightarrow C_1, \dots, C_n$, and $<$ is the relation of linear precedence determined by the grammar, then $\emptyset \in \Phi_r$ is *LP-Acceptable* if and only if whenever $\emptyset(C_i) < \emptyset(C_j)$, (i. e. precedes in the tree) there are no categories C_i', C_j' such that $\emptyset(C_i)$ extends C_i' , $\emptyset(C_j)$ extends C_j' , and ' $C_j' < C_i'$ ' is an LP-rule.

blems involving $\lceil +\text{FIRST} \rceil$ are resolved with the help of an EAIP statement.

(20) $\lceil +\text{FIRST} \rceil \text{ e}^{<<} \lceil \text{SUBCAT } kam \rceil$

(20) is a revised version of (17). This EAIP statement ensures that whenever *kam* occurs, one of its sister will be distinguished with the feature $\lceil +\text{FIRST} \rceil$ and will be the only constituent to linearly precede *kam*. In addition, by assuming a FSD (Feature Specification Default) of $\sim \lceil +\text{FIRST} \rceil$, this account guarantees that the feature $\lceil +\text{FIRST} \rceil$ is not instantiated unless required by LP statements or other grammaticality conditions such as (20). In addition, (20) requires that a $\lceil \text{SUBCAT } kam \rceil$ category cannot occur without a $\lceil +\text{FIRST} \rceil$ category as its sister. One consequence of this requirement is that a $\lceil +\text{FIRST} \rceil$ feature cannot be instantiated on a $\lceil \text{SUBCAT } kam \rceil$, since two $\lceil +\text{FIRST} \rceil$ sisters are disallowed by the LP statement (18). In this way, the problematic instantiation wrongly admitted by a simpler definition of IP-acceptable is ruled out. Thus we have shown that inadequacies in accounting for second position clitics with IP statements could be addressed when entailment annotation is introduced with the definition of EAIP-acceptable.

VI. Concluding Remarks

Cheng (1977: 156) observes that *kam* in Taiwanese is the form closest to Mandarin *ma*. *ma* is the unmarked form of Mandarin Yes-no question particle in the sense that it is modality-free and occurs in both assertive and rhetorical questions. *kam* shares this characteristic with *ma*. It is also worth pointing out that *kam* retains its sentential second position even though the majority of Taiwanese question particles, as described in Cheng (1977), are sentence-final. This article proposes a formal account of *kam* as a second-position sentential clitic in GPSG but offers no explanation of why a second-position clitic is an unmarked marker of interrogation in a language belonging to a family where sentence-final clitics are the norm. By showing that *kam* is phonologically dependent on the following element

with sandhi facts, we may safely conclude that it is a proclitic. Other liaison phenomena involving *kam*, however, have not been investigated in this article. Moreover, the picture presented in this paper is somewhat simplified. There are cases of *kam* occurring in positions which are possibly not second-positions. The only indisputable cases I know of involve modal auxiliaries such as *u* 'have' as in (21).²¹

- (21) *li tsang kam u lai*
you yesterday KAM have come
'Did you come yesterday?'

It is not immediately clear to me how sentences like (21) would be integrated into the formal account proposed in this paper. The fact that the syntax of Taiwanese is not well-studied forced me to leave many details unaccounted for. An adequate account of the cliticization of *kam* awaits more detailed study.

The present paper, however, does contribute to the formal study of second position clisis in general by proposing the entailment-annotated immediate precedence relation. This relation explicitly stipulates the co-occurrence of a second position clitic and the edge-marking feature [+FIRST]. The feature [+FIRST] is treated as a feature whose instantiation is marked and dependent on the interaction of other grammaticality conditions. Whether accounts along this line can be extended to second position clitics in other languages, and how the proposed EAIP statements affect the formal

21 Modality seems to be relevant, as suggested by the ungrammatical (ia), which lacks modal auxiliaries. The grammatical(ib), however, suggests that it is the presence of control verbs which makes the difference. This fact was pointed out to me by Ying-yu Sheu (p.c.).

(i) a. **i tsang kam lai*
s/he yesterday KAM come
b. *i miatsai kam bei/ai lai*
s/he tomorrow KAM want/need come
'Does s/he want/need to come tomorrow?'

properties of the grammar are two topics worth in-depth future studies.

* Research of this paper was supported by a National Science Council grant to the project 'Linguistic Survey of Chinese Dialects on Taiwan' (NSC 77-0301-H001-32RG). The comments of Ying-yu Sheu, and Drs. Ho Dah-an and Yang Siou-fang, which led to improvements in this paper, are gratefully acknowledged.

References

- Bresnan, Joan. 1982. *The Mental Representation of Grammatical Relations*. Cambridge: MIT Press.
- Chao, Yuen-Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Cheng, Robert L. 1977. Taiwanese Question Particles. *Journal of Chinese Linguistics*. 5.153-185.
- and Susie S. Cheng. 1977. *Phonological Structure and Romanization of Taiwanese Hokkien* [in Chinese]. Taipei: Students.
- Cheng, Susie S. 1981. *A Study of Taiwanese Adjectives*. Taipei: Students.
- Gazdar, Gerald, E. Klein, G. Pullum, and I. Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge: Harvard University Press. And Oxford: Blackwell.
- Huang, Chu-Ren. 1985. Chinese Sentential Clitics and Theories of Cliticization. Paper presented at the 1985 LSA Annual Meeting at Seattle.
- 1986. Coordination Schema and Chinese NP Coordination in GPSG. *Cahiers de Linguistique Asie Orientale*. XV. 1. 107-127.
- 1987. Mandarin Chinese NP *de*—A Comparative Study of Current Grammatical Theories. Cornell University Ph. D. Dissertation. Published as Special Publication of the Institute of History and Philology No. 93.
1989. Nankang: Academia Sinica.
- 1988. Cliticization and Type-shifting —A Unified Account of Mandarin NP *de*. [Revised first four chapters of Huang 1987] Bloomington: IULC.

- Kaisse, Ellen M. 1982. Sentential Clitics and Wackernagel's Law. *Proceedings of the First West Coast Conference on Formal Linguistics*. 1-14. Stanford: Stanford Linguistics Association.
- Kaplan, Ronald, and Joan Bresnan. 1982. *Lexical-Functional Grammar: A Formal System for Grammatical Representations*. Bresnan. 1982. 173-281. Cambridge: Massachusetts Institute of Technology Press.
- Klavans, Judith. 1982. *Some Problems in a Theory of Clitics*. Dissertation, University College London. [Reproduced by IULC.]
- 1985. The Independence of Syntax and Phonology in Cliticization. *Language*. 61.95-120.
- Lin, Shuang-fu. 1975. *The Grammar of Disjunctive Questions in Taiwanese*. Taipei: Students.
- Nevis, Joel A. 1985. A Syntactic Account of Second Position Clisis. S. Choi, D. Devitt, W. Janis, T. McCoy, and Z. -S. Zhang eds. *Proceedings of the Second Eastern States Conference on Linguistics*. 186-197. Columbus: The Ohio State University.
- 1986. Finnish Particle Clitics and General Clitic Theory. *The Ohio State University Working Papers in Linguistics* No. 33. Columbus: The Ohio State University.
- Sheu, Ying-yu. 1987. Chinese Morphosyntax. Paper presented at the Twentieth International Conference on Sino-Tibetan Languages and Linguistics. Vancouver, Canada.
- Shiu, Yu-Ling and Chu-Ren Huang. 1988. Unification-based Analysis and Parsing Strategy of Mandarin Particle Questions. *Proceedings of the International Computer Symposium*, 1988. pp. 38-43. Taipei: Tamkang University.
- Sun, Jackson T. -S. 1988. 安多若兒蓋藏語示證語氣初探, 'On Evidentiality in Amdo Tibetan Spoken at Ndzorge'. Mimeographed. Nankang: Institute of History and Philology, Academia Sinica.
- Zhu, De-Xi. 1985. 漢語方言裏的兩種反復問句, 'Two Patterns of A-not-A

- Questions in Chinese Dialects.' Zhongguo Yuwen 1985. 1. 10-20.
- Zwicky, Arnold. M. 1977. On Clitics. Bloomington: IULC.
- 1983. An Expanded View of Morphology in the Syntax-Phonology Interface. S. Hattori, and K. Inoue. Eds. Proceedings of the XIIIth International Congress of Linguistics. 198-208. Tokyo: CIPL.
- 1985. Clitics and Particles. Language. 61. 283-305.
- Zwicky, Arnold M., and Joel Nevis. 1986. Immediate Precedence in GPSG. Arnold M. Zwicky. Interfaces. (OSU Working Papers in Linguistics No. 32) 133-138. Columbus: Department of Linguistics, the Ohio State University.
- Zwicky, Arnold M., and Geoffrey K. Pullum. 1983. Cliticization vs. Inflection: English *n't*. Language. 59. 502-523.

摘 要

本文在概化詞組結構語法 (GPSG) 的架構下初步探討臺灣閩南語疑問助詞 *kam* (敢) 之語法與詞態現象；並論及此種現象在語法理論上的可能涵意。根據附加語 (clitic) 之定義，本文首先證明 *kam* 是個次位句附加語 (second-position sentential clitic)。在辨別 *kam* 和詞綴的分野時，其重要特徵包括了自由選擇其附主詞 (host word)，無特異詞態變化，不受語法律支配，與對其他詞綴的封閉性等。而 *kam* 與詞 (word) 的分野則在於其連音變調的現象，無代詞，嚴格詞序，及其固定位置等。*kam* 既非詞綴，亦非詞，只能是附加語。

本文藉 Zwicky 與 Nevis 二氏直接前行 (Immediate Precedence) 的觀念提出了 *kam* 的概化詞組結構語法分析，並首度提出了直接前行的規範定義。直接前行與共存前行 (由作者於 1985 提出) 之間的可能交互關係亦作了討論。最後，本文指出漢語與其他漢藏語系語言中末位句附加語最常見，閩南語中亦不例外。唯獨常用的疑問助詞 *kam* 是次位附加語，此一現象之歷史與語法解釋值得深究探討。